SHORT COMMUNICATION

MOLECULAR CHARACTERISATION OF *TOMATO YELLOW LEAF CURL VIRUS* Alm [Ma:Bk:02] IN MOROCCO: COMPLETE SEQUENCE AND GENOME ORGANISATION

N. Boukhatem[1], S. Jdaini[1], Y. Muhovsky[2], J.M. Jacquemin[2], C.L. Del Rincóne[3], M.J. Diéz[3] and A. Bouali[1*]

[1] *Laboratory of Genetic and Biotechnologies. Faculty of Sciences, University Mohamed 1st. P. O. Box 524, 60 000 Oujda, Morocco*
[2] *Centre Wallon de Recherches Agronomiques Gembloux, Département de Biotechnologie, 234, Chaussée de Charleroi, 5030 Gembloux, Belgium*
[3] *Universidad Polytechnica de Valencia, Camino de Vera, 46022 Valencia, Spain*

SUMMARY

Tomato yellow leaf curl disease is the main limiting factor for tomato (*Solanum lycopersicum*) production in Morocco. The PCR product obtained from a symptomatic sample was cloned and the complete DNA sequence (2781 bp) of the monopartite genome is reported of a *Tomato yellow leaf curl virus* isolate from north-east Morocco identified as TYLCV-Alm [Ma:Bk:02]. The circular genomic DNA contains six open reading frames encoding proteins of molecular weights ranging from 11 kDa to 40 kDa. Two proteins (V1 and V2) are located on the virion-sense strand whereas four other proteins (C1, C2, C3 and C4) are located on the complementary-sense strand. The two open reading frame groups are separated by an intergenic region of about 300 nucleotides, comprising a sequence capable of forming a stable hairpin loop structure containing the motif 5'-TAATATTAC-3'. The complete sequence obtained was closely related to that of *Tomato yellow leaf curl virus* (TYLCV) with nucleotide sequence identity ranging from 93% to 99%. Sequence comparison with other geminiviruses showed that the TYLCV isolate from north-eastern Morocco belongs to genus *Begomovirus* from the Old World, and is at least 99% identical to the sequence of Almeria TYLCV [TYLCV-(Alm)].

Key words: *Begomovirus*, TYLCV, TYLCV-Alm [Ma:Bk:02], PCR, cloning, DNA sequence.

Tomato yellow leaf curl disease (TYLCD) is one of the most devastating disorders of tomato crops in tropical and subtropical regions (Kheyr-Pour *et al*., 1991; Navot *et al*., 1991; Dry *et al*., 1993; Czosnek and Lalterrot, 1997). Causal agents are a group of members of the genus *Begomovirus* transmitted by the whitefly *Bemisia tabaci* Genn. Most begomoviruses have a bipartite genome consisting of two small circular single-stranded DNA (ssDNA) molecules (2.5-3.0 kb) (Rochester *et al*., 1994; Padidam *et al*., 1995). A few have a monopartite genome (2.8 kb) comprising six partially overlapping open reading frames (ORFs) (Kheyr-Pour *et al*., 1991; Noris *et al*., 1994; Crespi *et al*., 1995). Two ORFs in the virion sense (V1 and V2) and four in the complementary sense (C1, C2, C3 and C4) sequence are separated by an intergenic region (IR) that includes a sequence capable of forming a stable hairpin loop structure containing the motif 5'-TAATATTAC-3'. Such motif is common to all geminiviruses and the intergenic region is also the site of the origin of DNA replication (*ori*) (Stanley, 1995).

TYLCD symptoms were first described in 1964 in Israel, then detected in Italy, Spain, Portugal, Dominican Republic, Jamaica and Cuba (Polston and Anderson, 1997; Polston and McGovern, 1999). Viruses associated with TYLCD in these countries were identified as *Tomato yellow leaf curl virus* (TYLCV) and *Tomato yellow leaf curl Sardinia virus* (TYLCSV) by the International Committee on Taxonomy of Viruses (Stanley *et al*., 2005; Fauquet and Stanley, 2005).

In Morocco, TYLCD was first observed in 1996 in the region of El Jadida (west) and Berkane (north-east) (Jebbour and Abaha, 2002). The disease caused great yield losses in several tomato-growing provinces throughout the country. It was reported that 10% to 100% of tomatoes grown in these provinces were affected by the disease. The TYLCD-associated virus isolated from Morocco was classified as a member of the genus *Begomovirus* (Peterschmitt *et al*., 1999).

Understanding virus-vector-host interrelationships in the context of emerging problems can be achieved only by linking predicted evolutionary histories with epidemiology using molecular phylogenetic approaches. Identification and validation of informative molecular sequences are essential initial steps in this process. Here we report the complete nucleotide sequence and the genome organisation of a TYLCV isolate sampled in north-east of Morocco, which was identified as TYLCV-Alm [Ma:Bk:02]. Its nucleotide sequence was compared with other reported TYLCV isolates to determine identity rate and phylogenetic relationships.

Tomato plants (*Solanum lycopersicum*) showing begomovirus-like symptoms consisting of yellowing, leaf curling and distortion were collected from fields and

Corresponding author: A. Bouali
Fax: +212.36500603
E-mail: bouali@fso.ump.ma

greenhouses located in several areas in north-east Morocco (Berkane, Slimania, Ahfir). Infected leaves were stored in a freezer at -20°C.

Total DNA was extracted from collected leaf tissues according to Accotto *et al*. (2000), with minor modifications. PCR was done in a 25 µl reaction mixtures: 1 µl of 50 ng/µl DNA template, 1x Colorless GoTaq™ Buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$), 1.2 U of GoTaq™ DNA polymerase (Promega, Madison, WI, USA), 200 mM each of dNTPs, 2.5 mM $MgC1_2$, 300 nM of primers, and 0.8 µg BSA. The amplifications were carried out in an Icycler© Biorad thermocycler using the following programme: 4 min at 95°C followed by 35 amplification cycles (30 sec denaturation at 94°C, 40 sec annealing at 60°C, 1 min extension at 72°C) and a final extension at 72°C for 7 min. In every experiment a negative control with all PCR components except for the DNA template was included. Amplification products were analyzed by 1% agarose gel electrophoresis in 1% TAE buffer and stained with ethidium bromide with 1Kb smart ladder (Eurogentec, Seraing, Belgium) as a size marker. The gels were photographed under UV light.

Viral DNA extracted from an infected tomato plant was amplified using primer pairs MA13/MA26 (5'-AATGCAATCTTCGTCACC-3'/5'-CGCCCGTCTCGAAGGTTCG-3') and MA17/MA27 (5'-GAAAACATTTGTGAATCC-3'/5'-TGGAAATGATTATATCGCCTGGTCGC-3'), which yielded PCR products of the expected size (1292 bp and 1835 bp, respectively). Amplicons were cloned into pCR.2.1 TA vector (Invitrogen, San Diego, CA, USA) following the supplier's instructions, and transformed into competent *Escherichia coli* strain DH5α. The size of DNA inserts was checked by *EcoRI* restriction enzyme (Roche Diagnostics, Indianapolis, IN, USA) and DNA fragments were sequenced by Automatic Sequencer Global Edition IR_2 DNA Analyser. Four clones per amplicon were obtained from a minimum of two cloning experiments and sequenced in both directions using M13 universal primers.

DNA sequences of TYLCV-Alm [Ma:Bk:02] were analyzed using e-Seq (Li-COR e-Seq™ version 2) software, edited using the EditSeq Program (DNASTAR, Madison, WI, USA), then assembled and analysed again using DNAsis 2.5. Clustal W (MegAlign, DNASTAR) was used for multiple alignments of the viral genome sequence, which was compared with other reported geminivirus sequences available in GenBank. The genome organisation

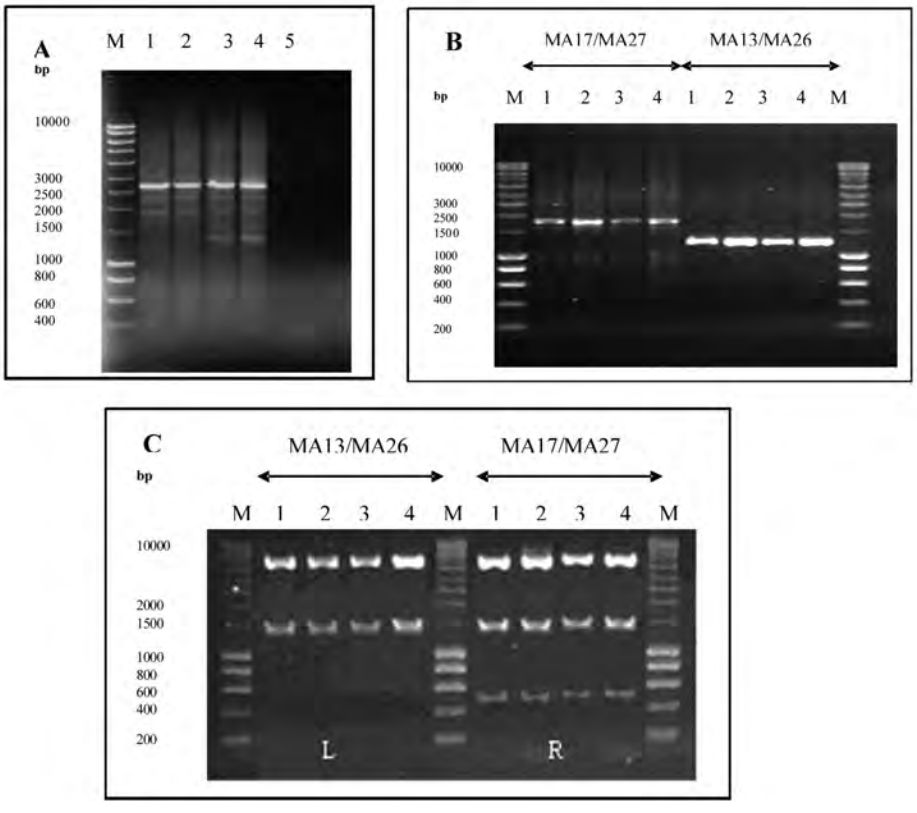

Fig. 1. Panel A: DNA amplified with primers MA26/MA27 using total DNA from four tomato plants infected by TYLCV, collected at Berkane (lanes 1-4); lane 5, healthy tomato (negative control). Panel B: PCR to detect TYLCV isolates from symptomatic field samples. Lane 1 to 4 correspond to TYLCV DNA amplified with MA17/MA27 using total DNA, prepared from sample 1 in panel A as template. Second set of lanes 1-4 corresponds to TYLCV DNA amplified with MA13/MA26 from the same sample as above. Panel C: restriction fragment length polymorphism analysis of clones obtained for TYLCV isolated from symptomatic field samples using EcoRI. Left (L) and right (R) panels represent clones of TYLCV DNA obtained with Ma13/Ma26 and MA17/MA27 respectively. In all panels lane M represents molecular weight markers.

of TYLCV-Alm [Ma:Bk:02] was determined and the functional ORFs on the virion sense and complementary sense were identified using Vector NTI software.

A virus-specific DNA fragment of 2781 bp was amplified by PCR using primers MA26/MA27 (Fig. 1, panel A), using as template total DNA extracted from an infected tomato plant. There was no amplification from total DNA extracted from healthy plants.

To determine the sequence of TYLCV-Alm [Ma:Bk:02] the two PCR products obtained using primers MA17/MA27 and MA13/MA26 (Fig. 1 panel B) were cloned into pCR.2.1 TA. Results from restriction enzyme digestion and sequencing showed that DNA fragments generated with MA17/MA27 contained a single *EcoRI* site (Fig. 1 Panel C). The sequences obtained using MA13/MA26 and MA17/MA27 primers were assembled, aligned and analysed using DNAsis 2.5 software. The full length nucleotide sequence was determined in both orientations.

The sense strand of the virus genome was identified and aligned with sequences available in GenBank. The entire genome of TYLCV-Alm [Ma:Bk:02] DNA was found to be 2781 bp long (accession No. EF060196).

The invariant TAATATTAC sequence located in the intergenic regions of geminiviruses was defined as the first nucleotides of the circular genome of TYLCV-Alm [Ma:Bk:02], whose organisation was very similar to that of monopartite begomoviruses. Six ORFs were found in TYLCV-Alm [Ma:Bk:02] genome, using the Genedoc software. The molecular weight of each protein was determined using Vector NTI software. The DNA sense strand encoded two overlapping ORFs with products of 29.9 kDa for ORF V1 (nt 315-1091) and 13.4 kDa for ORF V2 (nt 155-505). Amino acid sequences of V1 and V2 polypeptides and the position of these genes were similar to those of the coat protein and the movement protein of other geminiviruses infecting dicotyledonous hosts.

In the complementary sense orientation, four ORFs were found and designated C1 (40.6 kDa), C2 (15.4 kDa), C3 (16.0 kDa), and C4 (11.0 kDa) based on size and sequence similarities with polypeptides of other geminiviruses (Table 1). The position of ORF C1 (nt 2622-1549), C2 (nt 1640-1233), C3 (nt 1492-1088) and C4 (nt 2471-2178) was determined and found to be comparable to that of the corresponding genes of other geminiviruses (Noris *et al*., 1994; Gafni, 2003). An intergenic region located between 2623 nt and 154 nt, comprised a sequence capable of forming a stable hairpin loop structure containing the invariant sequence 5'-TAATATTAC-3' and the initiation site of the rolling circle DNA replication (*ori*) of geminiviruses (Arguello-Astorga *et al*., 1994; Stanley, 1995).

The complete nucleotide sequence of TYLCV-Alm [Ma:Bk:02] was compared with 20 other geminivirus sequences from GenBank or EMBL databases. In particular, the putative products of TYLCV-Alm [Ma:Bk:02] ORFs were compared with the corresponding proteins of several TYLCV isolates in a pairwise fashion to determine the percentage of identity and the relatedness of the sequences within the group (Table 1). ORF V1 (CP) sequence found in similar position in all Old World geminiviruses overlapped the amino terminus of ORF V2 (movement protein). TYLCV-Alm [Ma:Bk:02] showed the highest percentage identity (98.8%) in the ORF V1 with TYLCV (X15656), TYLCV-Mld (X76319), TYLCV-[Alm] (AJ489258) and Tomato yellow leaf curl virus [Cuba] (AJ223505). Comparison of C1, C2, C3 and C4 sequences of our isolate shared the highest identity with TYLCV reported from Almeria (Table 1).

Sequence of cloned PCR fragments spanning the complete TYLCV-Alm [Ma:Bk:02] DNA (2781 bp) exhibited the typical genome organisation of begomoviruses originating from the Old World, with six conserved ORFs. The first complete sequences of TYLCVs were reported in 1991 for isolates from Sardinia (TYLCSV) and Israel (TYLCV) (Kheyr-Pour *et al*.,

Table 1. Percentage of nucleotide identity of total DNA, and percentage of amino acid identity of the putative proteins of TYLCV and TYLCSV isolates from different countries.

	TYLCV					TYLCSV	
	Mild	Israel	Cuba	Dominican Republic	Almeria	Italy	Morocco
Total DNA identity (nt%)							
	93.7	98	98.1	98.1	99.2	76.2	75.7
Amino acid identity (aa%)							
V1 (CP)	98.8	96.9	98.8	98.4	98.8	88.8	88
V2 (MP)	99.1	99.1	99.1	99.1	99.1	81.0	81.0
C1 (Rep)	89.0	96.9	96.9	76.1	98.2	77.3	96.7
C2	95.6	96.3	96.3	96.3	97.8	64.4	64.4
C3	95.5	97.0	95.5	95.5	97.8	63.4	65.7
C4	52.7	94.8	96.9	96.9	95.9	49.5	43.0

1991; Navot *et al*., 1991), whereas the first complete sequence of a Moroccan isolate was reported in 2004 (accession No. AY702650; EL Mehrach *et al*., 2007) and showed the highest identity (99%) with the TYLCV isolates from Spain (TYLCSV-[ES1]; Z25751).

Much genetic information of partial sequences of V1 and V2 was obtained for TYLCV isolates from various locations in Morocco (Berkane, Meknes and Casablanca) (Peterschmitt *et al*., 1999; Jdaini *et al*., 2005; Tahiri *et al*., 2007) (Table 1). Based on sequence comparisons, TYLCV-Alm [Ma:Bk:02] was most closely related to TYLCV-[Alm] with 99% nucleotide identity. A similar result was obtained to the deduced amino acid sequences of the putative ORF products (Table 1).

It is worth noting that even if the two completely sequenced TYLCV isolates from Morocco (Berkane and Agadir) belong to two different species, i.e. TYLCV and TYLCSV, they and have the highest sequence identity with the Spanish isolates from Almeria and Murcia, respectively. This finding supports the hypothesis that the disease was introduced to Morocco from Spain, most probably via infected tomato seedlings or whitefly.

To our knowledge, this is the first time that an isolate of the type strain of TYLCV is reported from Morocco, which differs from other TYLCV isolates previously characterised in this country, that grouped with mild strains of TYLCV (Peterschmitt *et al*., 1999; Tahiri *et al*., 2007).

ACKNOWLEDGEMENTS

The authors thank Z. Chafik for the assistance in collecting samples, and Dr. A. Kharraki for critical reading of the manuscript. The work was supported by PROTARS grant from the Moroccan Ministry of Education and the C.U.D., Belgium.

REFERENCES

Accotto G.P., Navas-Castillo J., Noris E., Moriones E., Louro D., 2000. Typing of tomato yellow leaf curl viruses in Europe. *European Journal of Plant Pathology* **106**: 179-186.

Arguello-Astorga G.R., Guevara-Gonzalez R.G., Herrera-Estrella L.R., Rivera-Bustamante R.F., 1994. Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. *Virology* **203**: 90-100.

Crespi S.E., Noris A.M., Vaira A.M., Accotto G.P., 1995. Molecular characterization of cloned DNA from tomato yellow leaf curl virus isolate from Sicily. *Phytopathologia Mediterranea* **34**: 93-99.

Czosnek H., Laterrot H., 1997. A worldwide survey of tomato yellow leaf curl viruses. *Archives of Virology* **124**: 1391-1406.

Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A., 1993. Nucleotide sequence and genome organization of tomato leaf curl geminivirus. *Journal of General Virology* **74**: 147-151.

El Mehrach K., Sedegui M., Hatimi H., Tahrouch S., Arifi A., Czosnek H., Nakhla M.K., Maxwell D.P., 2007. Molecular characterization of a Moroccan isolate of Tomato yellow leaf curl Sardinia virus and differentiation of the Tomato yellow leaf curl virus complex by the polymerase chain reaction. *Phytopathologia Mediterranea* **46**: 185-194.

Fauquet C.M., Stanley J., 2005. Revising the way we conceive and name viruses below the species level: A review of geminivirus taxonomy calls for new standardized isolate descriptors. *Archives of Virology* **150**: 2151-2179.

Gafni Y., 2003. Tomato yellow leaf curl virus the intracellular dynamics of a plant DNA virus. *Molecular Plant Pathology* **4**: 9-15.

Jdaini S., Boukhatem N., Bouali A., Diez Niclos M., Rincon C.L., 2005. Caractérisation moléculaire du virus TYLC de la région de Berkane. *Comptes Rendues des Journées Nationales de Microbiologie, Oujda 2005*: 19.

Jebbour F., Abaha L., 2002. Situation de la maladie des feuilles jaunes en cuillère de la tomate (TYLCV) au Maroc. *Bulletin OEPP/EPPO Bulletin* **32**: 17-19.

Kheyr-Pour A., Bendahmane M., Matzeit N., Accotto G.P., Crespi S., Gronenborn B., 1991. Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus. *Nucleic Acids Research* **19**: 6763-6769.

Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H., 1991. Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single genomic component. *Virology* **185**:151-161.

Noris E., Hidalgo E., Accotto G.P., Moriones E., 1994. High similarity among the tomato yellow leaf curl virus isolates from the West Mediterranean Basin: the nucleotide sequence of an infectious clone from Spain. *Archives of Virology* **135**: 165-170.

Padidam M., Beachy R.N., Fauquet C.M., 1995. Classification and identification of geminiviruses using sequence comparaisons. *Journal of General Virology* **76**: 249-263.

Peterschmitt M., Granier M., Aboulama S., 1999. First report of Tomato yellow leaf curl geminivirus in Morocco. *Plant Disease* **83**: 1074.

Polston J.E., Anderson P.K., 1997. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. *Plant Disease* **81**: 1358-1369.

Polston J.E., McGovern R.J., 1999. Introduction of tomato yellow leaf curl virus in Florida and implications for the spread of this and other geminiviruses of tomato. *Plant Disease* **83**: 984-988.

Rochester D.E., DePaulo J.J., Fauquet C.M., Beachy R.N., 1994. Complete nucleotide sequence of the geminivirus tomato yellow leaf curl virus Thailand isolate. *Journal of General Virology* **75**: 477-485.

Stanley J., 1995. Analysis of African cassava mosaic virus recombinants suggests strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. *Virology* **206**: 707-712.

Stanley J., Bisaro D.M., Briddon R.W., Brown J.K., Fauquet C.M., Harrison B.D., Rybicki E.P., Stenger D.C., 2005. Family *Geminiviridae*. In: Fauquet C.M., Mayo M.A., Maniloff J., Desselberger U., Ball L.A. (eds), Virus taxonomy, Eight Report of the International Committee on Taxonomy of Viruses, pp. 301-326. Elsevier/Academic Press, London, UK.

Tahiri A., Sekkat A., Bennani A., Granier M., Delvare G., Peterschmitt M., 2007. Distribution of tomato-infecting begomoviruses and *Bemisia tabaci* biotypes in Morocco. *Annals of Applied Biology* **149**: 175-186.

Received June 6, 2007
Accepted October 1, 2007